Network Science **8** (S1): S4–S25, 2020
doi:10.1017/nws.2020.3

CAMBRIDGE
UNIVERSITY PRESS

ORIGINAL ARTICLE

A model for the dynamics of face-to-face interactions in social groups

Marion Hoffman[1*], Per Block[2], Timon Elmer[1] and Christoph Stadtfeld[1]

[1]Chair of Social Networks, ETH Zürich, Zürich, Switzerland (e-mails: timon.elmer@gess.ethz.ch, christoph.stadtfeld@ethz.ch) and [2]Department of Sociology, University of Oxford, Oxford OX1 1JD, UK (e-mail: per.block@sociology.ox.ac.uk)
*Corresponding author. Email: marion.hoffman@gess.ethz.ch

Special Issue Editors: Hocine Cherifi, Stanley Wasserman

Abstract

Face-to-face interactions in social groups are a central aspect of human social lives. Although the composition of such groups has received ample attention in various fields—e.g., sociology, social psychology, management, and educational science—their micro-level dynamics are rarely analyzed empirically. In this article, we present a new statistical network model (DyNAM-i) that can represent the dynamics of conversation groups and interpersonal interaction in different social contexts. Taking an actor-oriented perspective, this model can be applied to test how individuals' interaction patterns differ and how they choose and change their interaction groups. It moves beyond dyadic interaction mechanisms and translates central social network mechanisms—such as homophily, transitivity, and popularity—to the context of interactions in group settings. The utility and practical applicability of the new model are illustrated in two social network studies that investigate face-to-face interactions in a small party and an office setting.

Keywords: Network models; social interactions; group interactions; statistical models; social sensors; relational events; actor-oriented models.

1. Introduction

The social occasions when people gather around particular events, such as weddings, dinner parties, joint activities, lectures, or conferences, mark the times and places when they escape their routine everyday lives and interact with each other. Here, experiences are shared, common identities confirmed, informal hierarchies negotiated, and collective memories created (A. Goffman, 2019; E. Goffman, 1967; Simmel & Hughes, 1949; Wynn, 2016). Unsurprisingly, individuals invest a considerable amount of time and resources into joining or organizing social gatherings. Some social occasions even mark turning points in the course of individuals' personal and professional lives (Goffman, 2019).

The social occasion has a further unique role in defining individuals' social lives, as it is the birthplace of many social relations that have a crucial and lasting impact on their life courses. For example, family routines of repeated interactions, such as shared mealtimes, are paramount in the development of emotional bonds between parents and their children (Spagnola & Fiese, 2007). Outside of the home, schoolyard interactions and extracurricular activities play a crucial role in the organization of children's social groups (Moody, 2001). In the professional lives of adults, meetings within formal settings such as offices and conferences become instrumental in providing individuals the means to achieve their professional goals. For example, exchanges during international conferences might promote better collaboration between scientists (Wang et al., 2017) or facilitate political negotiations between different countries representatives (Schroeder & Lovell, 2012). Smaller gatherings such as coffee breaks or meetings in an office allow employees of the same organization to share knowledge and support each other (Lazega et al., 2006; Wineman et al., 2014). Such interactions eventually contribute to defining families' identities (Spagnola & Fiese, 2007), social groups composition (Lazarsfeld & Merton, 1954; Moody, 2001), or organizational structures (Allen et al., 2007; Lazega et al., 2006). Following its importance for understanding human relationships, this paper proposes a statistical framework to analyze dynamic group interactions occurring during social occasions.

1.1 Studying group-based interactions

The importance of group-based gatherings and interactions is a consistent theme in the literature on network formation. Numerous social network theories underline the influence of social occasions on social organization (Carley & Krackhardt, 1996; Fischer, 1982; Lazarsfeld & Merton, 1954; Lazega et al., 2006; Moody, 2001) and often rely, implicitly or explicitly, on the occurrence of face-to-face interactions to explain the formation of social ties. Social occasions are also crucial for the maintenance of ties, as people even meet most of their existing social contacts only in group-based settings. Dyadic social interactions are indeed often reserved for their innermost social circles, composed of romantic partners, family, and close friends (Block, 2018; Fischer, 1982). Umbrella concepts for such group-based meetings that only differ in nuances are social circles (Simmel & Hughes, 1949), social foci (Feld, 1982), social settings (Pattison & Robins, 2002), or social situations (Block, 2018).

While such approaches acknowledge the importance of interactions within social settings of individuals, they implicitly assume that everybody within this setting interacts with all others equally. Thus, the further structuring of interactions within these occasions is not analyzed. However, social interactions are driven by two complementary processes: a "meeting" process that brings different individuals together within specific occasions and a "mating" process that defines how these individuals will interact within these occasions (Verbrugge, 1977). Applied to our case, most network research considers individuals "meeting" within the same social occasions but disregards the "mating" dimension of these situations. In our model, we explicitly model the "mating" dimension; that is, we aim to analyze how individuals move between interactions within social gatherings. By doing so, we can examine whether individual interaction dynamics have similar patterns as those observed in more durable ties, such as homophily or transitivity.

Interactions within groups have nonetheless received scholarly attention. Micro-sociological and social psychological theories have stressed the importance of social processes occurring during face-to-face interactions. On the one hand, theories by Goffman (1967) or Collins (2014) on face-to-face interactions and interaction rituals have defined a wide range of social behaviors as the elementary bricks of social interactions. For example, Goffmans theory explores the complex interdependencies between basic signals such as glances or gestures, as well as more complex actions, such as adaptation to conversational norms or turn-taking, among different individuals interacting in a group.

On the other hand, social psychologists have emphasized the idea that some social processes might result from "group dynamics" (as coined by Lewin, 1947). Such processes can help us understand how social interactions in and between social groups lead to individual and group-level outcomes. Most notably, Tajfel's social identity theory and the minimal group paradigm (Tajfel et al., 1979; Tajfel, 1970) posit that individuals swiftly categorize themselves and others in such situations into social groups that guide their behavior toward in and out-group members. Other influential approaches focus more on how individuals engage in social interactions with groups, such as Bales (1950)'s interaction process analysis. In these approaches, not only the social dynamics of groups are of interest but also how individual characteristics affect the formation of social interactions in groups.

1.2 Motivation for a new model

Although interest in social interactions within group-based settings is well established, researchers have long-lacked methods to collect high-resolution data on such interactions. Recent advances in current sensor technologies, such as Bluetooth or RFID badges (Cattuto et al., 2010; Elmer et al., 2019), sociometric badges (Olguín et al., 2008; Pentland, 2008), or Wi-Fi traces (Hong et al., 2016; Sapiezynski et al., 2015), now provide new opportunities to measure the occurrence as well as the composition and evolution of face-to-face interactions. However, previous analyses of these data have rarely taken the group nature of face-to-face interactions into account, i.e., that most conversations and interactions at social occasions happen in multi-person cliques, with some rare exceptions (Sekara et al., 2016). Moreover, most network methods consider interactions as dyadic events rather than dynamic "multi-person ties".

The current paper addresses this gap by providing a statistical model that can express the social mechanisms underlying face-to-face interactions between two or more individuals in larger social settings. Drawing from relational event models (Butts, 2008; Stadtfeld, 2012), and in particular from the Dynamic Network Actor Model (DyNAM, Stadtfeld & Block, 2017; Stadtfeld et al., 2017), we propose a model that can take into account the particular format and the interdependencies of such data. These include interdependencies between individuals (i.e., interactions between individuals depend on the interaction patterns of other individuals), between time points (i.e., current interactions depend on previous interactions), and between group members (i.e., interactions of an individual in a group depend on the group's composition). To this end, we extend the definitions of classic mechanisms commonly observed in social networks based on pairwise ties to the group case. For instance, homophily and transitivity are relevant in group-based interactions but require elaboration beyond dyadic or triadic configurations.

The remainder of this article is structured as follows: Section 2 defines the statistical model; Section 3 discusses the implementation of theoretical mechanisms. Estimation and model implementation are outlined in Section 4. We present an example application to a data set of social interactions collected via video recordings during a social gathering in Section 5 and a larger case study on daily interactions in an office setting in Section 6. The article concludes with a brief discussion of the model in relation to the literature and future research avenues.

2. Statistical model

2.1 Previous models

The method presented here builds upon a rich tradition of statistical frameworks that aim at explaining the emergence and dynamics of social networks (for a review, see Robins, 2015). Modeling social network ties as dependent variables require to pay particular attention to dependencies among observations when examining network mechanisms such as reciprocity, homophily, transitivity, or preferential attachment. Frequently applied statistical network models are Exponential Random Graph Models (ERGMs, Lusher et al., 2013) and Stochastic Actor-Oriented Models (SAOMs; Snijders, 1996; Snijders et al., 2010). They explicitly express complex dependencies between individual and tie variables to draw inference about relations measured at one or multiple points in time (e.g., friendship and collaboration).

More recently, statistical methods for the analysis of *relational events* (Butts, 2008) have been proposed. Relational events are sequences of discrete dyadic events, such as phone calls, social media activities, or contracts for which exact time-stamped information can be measured. The first of these models is the Relational Event Model (REM; Butts, 2008) which expresses event sequences as continuous-time Markov chains, in the tradition of earlier approaches in social network literature (e.g., Holland & Leinhardt, 1977; Snijders, 1996; Wasserman, 1980).

A number of subsequent models extend the REM (e.g., Brandes et al., 2009; Stadtfeld & Geyer-Schulz, 2011; Perry & Wolfe, 2013; Marcum & Butts, 2015; Vu et al., 2015; Schecter et al., 2018;

Amati et al., 2019; Brandenberger, 2019; Lerner & Lomi, 2019; Mulder & Leenders, 2019). The model introduced in this paper extends the Dynamic Network Actor Model (DyNAM; Stadtfeld & Block, 2017; Stadtfeld et al., 2017) that formulates an actor-oriented framework for relational events. In the DyNAM framework, an event is decomposed in first an actor's decision to send an event at a certain point in time, and second the decision to send this event to a specific actor. These two steps define two conditionally independent models called the rate and the choice models. The new model extends this logic to model face-to-face interactions.

2.2 General logic

Face-to-face interactions are a specific case of relational events that can not only occur dyadically but also between more than two people. When we speak of groups in this article, we refer to two or more individuals who are part of the same face-to-face conversation. The first assumption of this model is that individuals can only be part of one or no group at a time (i.e., membership in groups is exclusive). A second assumption is that changes in group compositions can be reasonably approximated by sequences of non-simultaneous individual decisions to join and leave groups. An interaction group forms when at least two individuals meet; actors might join or leave this group, but it will exist as long as at least two individuals keep interacting.

Conceptually, we can think of group interactions as a two-mode network with a dynamically changing node set, where the first mode refers to individuals or actors and the second mode to groups. In our mathematical definitions, the second mode comprises nodes representing interaction groups (i.e., group nodes) as well as nodes representing isolated actors (i.e., isolate nodes). At any given point in time, each actor is thus affiliated with exactly one node in the second mode. A new group emerges when one actor decides to start an interaction with another isolated actor. In this case, these two actors belong to the same group node. While this group node exists, other isolated actors can similarly join the group. The group dissolves when only two actors belong to it, and one of them decides to leave. From there, the two actors are only affiliated with an isolate node.

This paper proposes a model that extends the DyNAM framework to explain sequences of joining and leaving events in the context of group interactions. We call this model DyNAM-i, the letter i referring to interaction. Following an actor-oriented logic, the model consists of three steps:

1. The first step describes how individuals who are not part of an interaction (isolates) decide to join a group or another isolate.
2. The second step describes how individuals who decided to approach a group (step 1) choose which group or isolate to join.
3. The third step describes how individuals who are part of a group decide to leave the interaction (and become isolates).

This formulation differs from the original DyNAM definition in three ways. First, a third step is added to model the end of an interaction, which allows us to explain both its occurrence and its duration. Second, the sets of actors who can make decisions in steps 1 and 3 are restricted by the actors' position in the two-mode network (e.g., only isolates can join a group, only individuals in a group can leave their group). Third, the mechanisms dictating the choices of actors in step 2 are no longer defined at a dyadic level, which will become apparent in Section 3.

The model aims at explaining sequences of "joining events" and "leaving events" between actors and groups. A joining event occurs when an actor starts interacting with another actor (hence creating a new group) or an already existing group of actors. It is defined by the sequence of steps 1 and 2 of the model. A leaving event occurs when an actor leaves the group she currently belongs to, which is determined by step 3.

By modeling the three steps separately, we can test a wide range of hypotheses from social theories related to individuals' preferences given available interaction opportunities. Social mechanisms can consider individual attributes, group attributes, other social ties (e.g., friendship ties), and prior changes. By modeling step 1, we can examine how individual characteristics and prior activities predict individuals' tendencies to join groups. One can ask, for example, if personality traits like extraversion explain the preference of not being isolated in the interaction network. Modeling step 2 allows us to understand how individual characteristics, characteristics of the group and its members, and prior activities affect how individuals choose a group to join. For instance, some individuals might prefer to interact in groups with many members of the same gender (individual-group homophily). Finally, the model for step 3 allows us to test whether individual characteristics and group characteristics explain the tendency of individuals to stay in a conversation or to leave. One could then examine, for example, whether people are more likely to leave a group interaction in which they have no friends.

2.3 Notation

For the mathematical formulation of the model, we propose the following notation. We consider

$$\mathscr{I}(t) = \{1, ..., n(t)\}$$

the set of actors who are present at time t and

$$\mathscr{G}(t) = \{1, ..., m(t)\}$$

the set of interaction groups at time t. For better readability, we may omit the indicator t in some notations below. The *interaction network* at a given time is denoted by an $n(t) \times m(t)$ matrix $X(t)$, whose elements $x_{i,g}(t)$ equal to one if the actor i is interacting within the group g at time t, and 0 otherwise:

$$X(t) = [x_{i,h}]_{i \in \mathscr{I}(t), h \in \mathscr{G}(t)}.$$

$X(t)$ is a two-mode network. At each point in time, we define for $\mathscr{I}(t)$ the subset $\mathscr{I}^{(0)}(t)$ that contains all isolated actors (those who are not part of a group interaction) and the subsets $\mathscr{I}^{(g)}(t)$ for actors who are members of the groups $g \in \mathscr{G}(t)$. The set $\mathscr{I}(t)$ is, therefore, the (disjoint) unions of these sets:

$$\mathscr{I}(t) = \bigcup_{0 \leq h \leq m(t)} \mathscr{I}^{(k)}(t),$$

$$i \in \mathscr{I}^{(0)}(t) \Leftrightarrow \forall h \in \mathscr{G}(t) : x_{i,h} = 0$$

$$i \in \mathscr{I}^{(g)} \Leftrightarrow x_{i,g} = 1$$

$$\forall h, h' : I^{(h)} \cap I^{(h')} = \{\}, \text{if } h \neq h'$$

In order to account for the history of events, we also define the *network of past interactions* as an $n(t) \times n(t)$ matrix $X^{\text{past}}(t)$, with $x_{i,j}^{\text{past}}(t)$ indicating the number of interactions in which both actors i and j participated before time t. Additional networks of past interactions may express past interactions that occurred within a specific time window δ and include all interactions in the interval $[t - \delta, t]$.

$$X^{\text{past}}(t) = [x_{i,j}^{\text{past}}]_{i,j \in \mathscr{I}(t)}$$

Each actor can be characterized by a set of p stable or changing attributes (e.g., gender, personality traits, current mood) that define the *actor attribute matrix*

$$A(t) = [a_{i,h}]_{i \in \mathscr{I}(t), 1 \leq h \leq p}$$

containing all p attributes of actors at time t. Similarly, a group attribute matrix can be defined to represent the context in which groups form or the content of the interactions within them.

Additional relational information relevant to the analysis, such as other interpersonal ties between actors, can be defined as $n(t) \times n(t)$ matrices denoted by:

$$Z^{(1)}(t), Z^{(2)}(t), \ldots$$

The dependent variable of our model consists of a sequence Ω of *joining* and *leaving events* between actors and interaction groups. An event ω is defined as the quadruplet

$$\omega = (t_\omega, i_\omega, g_\omega, d_\omega)$$

with t_ω being the time at which the event occurs, i_ω the actor initiating the event, g_ω the group that i is joining or leaving, and d_ω the description of the event, i.e., a boolean indicating whether ω_i is a joining event ($d_\omega = 1$) or leaving event ($d_\omega = 0$). The dependent variable explained by the model is the complete, ordered sequence of all observed events in the data:

$$\Omega = \{\omega_k\}_k.$$

2.4 Model definition

In order to define a probability model, we consider a continuous-time Markov chain similar to the one presented by Stadtfeld & Block (2017). This Markov process operates on the state space y that represents all possible situations that can exist among a given set of actors. To fulfill the Markov criterion, the process state $y(t)$ must contain all information relevant for modeling the probability of an event to occur at time t. It is defined as a discrete vector:

$$y(t) = \left(\mathscr{I}(t), \mathscr{G}(t), X(t), X^{\text{past}}(t), A(t), B(t), Z^{(1)}(t), Z^{(2)}(t), \ldots\right) \tag{1}$$

It should contain information on the individuals present, the existing groups, and the current group partition. Depending on the mechanisms of interest, it can also include information on past interactions, individual or group attributes, and relational ties.

The stochastic process is entirely defined by the transition rate matrix Q whose elements $q(y_1 \to y_2)$ describe the rate of the transitions from a state y_1 to another state y_2 and therefore specify the occurrence and timing of interaction events. The use of a Markov process simplifies the definition of the probability model for events ω to occur and make the estimation of the model parameters tractable.

A graphical representation of the modeled process is shown in Figure 1. It follows the conceptual three-step logic outlined above. We define τ_i^{joining} and τ_i^{leaving} as Poisson rates for joining a group or an isolate (step 1) and leaving a group (step 3), respectively. Step 1 and 3 will be referred to as the *rate model* (green box in Figure 1). We define $p_{i,g}$ as the probability for an actor i to choose a group g among all options available (step 2). We will refer to this as the *choice model* (blue box in Figure 1).

For a given event ω, with the boolean d_ω indicating whether it is a joining or leaving event, let the previous process state be y^0, and y^ω the process state once updated according to the event ω. The transition rate from y^0 to y^ω is defined by a combination of the three sub parts:

$$q(y^0 \to y^\omega) = d_\omega \tau_{i_\omega}^{\text{joining}} p_{i_\omega, g_\omega} + (1 - d_\omega)\tau_{i_\omega, g_\omega}^{\text{leaving}} \tag{2}$$

This combination of three sub processes makes specific assumptions. The timing of joining events (step 1) is assumed to be conditionally independent of the actors' choices (step 2). The two timing processes (step 1 and step 3) are also assumed to be conditionally independent given the process state y^0. Thereby, we can define the process as a composite Poisson process, in which waiting times before joining or leaving an interaction follow two Poisson processes and the choice of interaction groups follows a discrete probability model.

https://doi.org/10.1017/nws.2020.3 Published online by Cambridge University Press

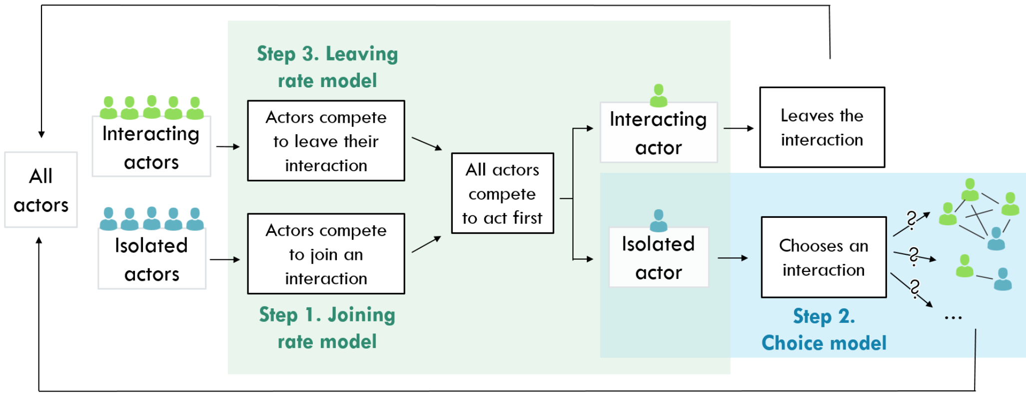

Figure 1. Description of the three-step process modeled by DyNAM-i.

Step 1 of the model is expressed as a Poisson process. The rates τ_i^{joining} are defined for each isolated actor $i \in \mathscr{I}^{(0)}$. Different actors can have different tendencies to join interactions, and these tendencies can evolve through time. The joining rate is mathematically expressed as an exponential link function for a linear combination of a statistics vector $(s_{k,i}^{\text{joining}})_k$, conditioned by the parameter vector $(\alpha_k^{\text{joining}})_k$, where k refers to the index of the effects with which the model was specified. Parameter $\alpha_0^{\text{joining}}$ is the intercept of the model and expresses the baseline waiting time for joining a group when all other statistics are null.

$$\tau_i^{\text{joining}}(y(t)|\alpha^{\text{joining}}) = \exp\left(\alpha_0^{\text{joining}} + \sum_k \alpha_k^{\text{joining}} s_{k,i}^{\text{joining}}(y(t))\right) \tag{3}$$

Step 3 of the model is expressed as a Poisson process in a similar way. The rates $\tau_{i,g}^{\text{leaving}}$ are defined for each actor $i \notin \mathscr{I}^{(0)}$ that is currently interacting in a group g. It is also expressed as a linear combination of a statistic vector $(s_{k,i,g}^{\text{leaving}})_k$ conditioned by the parameter vector $(\alpha_k^{\text{leaving}})_k$ that includes another intercept $\alpha_0^{\text{leaving}}$. Here, the statistics not only depend on the focal actor but also on the characteristics of the group.

$$\tau_i^{\text{leaving}}(y(t)|\alpha^{\text{leaving}}) = \exp\left(\alpha_0^{\text{leaving}} + \sum_k \alpha_k^{\text{leaving}} s_{k,i,g}^{\text{leaving}}(y(t))\right) \tag{4}$$

These two Poisson processes are competing. The actions of an actor, however, only follow one of the two rate functions, depending on whether she is isolated or not. The rate of all actors is denoted by:

$$\tau_i(y(t),\alpha) = \begin{cases} \tau_i^{\text{joining}}(y(t)|\alpha^{\text{joining}}),\ i \in \mathscr{I}^{(0)} \\ \tau_i^{\text{leaving}}(y(t)|\alpha^{\text{leaving}}),\ i \notin \mathscr{I}^{(0)} \end{cases} \tag{5}$$

with $\alpha = (\alpha^{\text{leaving}}, \alpha^{\text{joining}})$ the parameter vector for the whole rate model. The probability of the next (joining or leaving) event to be initiated by the actor i can be calculated as:

$$p(\text{actor } i \text{ is active next}|y(t),\alpha) = \frac{\tau_i(y(t)|\alpha)}{\sum_{j\in\mathscr{I}(t)}(y(t)|\alpha)} \tag{6}$$

Step 2 of the model is expressed as a multinomial choice probability (McFadden, 1974), similarly to DyNAM and SAOM frameworks. Provided that an actor decided to join an interaction (step 1), she tries to optimize an objective function defined for each interaction group that is currently present. This objective function is defined by a linear combination of a statistic vector $(s^{\text{choice}}_{k,i,g})_k$ with k elements, conditioned by a parameter vector $(\beta_k)_k$:

$$\mu_{i,g}(y(t)) = \exp\left(\sum_k \beta_k s^{\text{choice}}_{k,i,g}(y(t))\right) \tag{7}$$

The probability of actor i choosing the group g is expressed as the ratio of the objective function $\mu_{i,g}$ to the sum of all objective functions for available options. In other words, the actor compares all her options according to Equation (7).

$$p(g|i, t, \beta) = \frac{\mu_{i,g}(y, t)}{\sum_{h=1}^{m(t)} \mu_{i,h}(y, t)} \tag{8}$$

We have translated the three conceptual steps of the model into a mathematical framework. The rates and the choice function are conditional on parameters α and β that are unknown but are estimated from the data. The probability of the data given the parameters (the likelihood) can be expressed similarly to the formulations derived in Stadtfeld (2012) and in Stadtfeld & Block (2017). Section 3 provides possible specifications for the statistics $(s^{\text{joining}}_{k,i})_k$, $(s^{\text{leaving}}_{k,i,g})_k$, and $(s^{\text{choice}}_{k,i,g})_k$. Section 4 details the implementation of the maximum likelihood estimation routine.

3. Model specifications

The statistics $(s^{\text{joining}}_{k,i})_k$, $(s^{\text{leaving}}_{k,i,g})_k$, and $(s^{\text{choice}}_{k,i,g})_k$ introduced in Equations (3), (4), and (7) define the characteristics of the observed events that influence the probability of their occurrence. Prominent examples of such network mechanisms are reciprocity, transitivity, homophily, and preferential attachment (Robins, 2015). Mathematical specifications of such network mechanisms have been specified for statistical models such as ERGMs, SAOMs, REMs, and DyNAMs (Butts, 2008; Lusher et al., 2013; Snijders, 2017; Stadtfeld, 2012). Some of these mechanisms can be translated into our model, while others need to be adapted to the notion of group interactions. In the following, we discuss three classes of network mechanisms that relate to (i) actor or group attributes, (ii) relational ties between actors in other networks, and (iii) previous interaction events. The naming conventions resemble those of previous instances of DyNAM models (Stadtfeld & Block, 2017; Stadtfeld et al., 2017).

3.1 Mechanisms related to actor and group attributes

One class of mechanisms is concerned with how individual attributes affect the tendency to join or leave groups (steps 1 and 3) and how group attributes affect individuals' tendency to choose or leave a group interaction (steps 2 and 3).

Snijders & Lomi (2019) classify attribute-related network mechanisms into four classes of effects, namely *sociability* (i.e., the tendency for actors scoring high on a particular attribute to send more ties), *aspiration* (i.e., the tendency to bond with individuals scoring high on a certain attribute), *homophily* or *assortativity* (i.e., the tendency to bond with similar others), or *conformity* (i.e., the tendency to bond with individuals who are closer to a certain norm).

Following the idea of *sociability*, we define the *ego* effect that captures the influence of an individual's attribute on her tendency to join or leave interactions in steps 1 and 3. One might, for example, test whether introverted individuals are less likely to interact with others or to stay in an

https://doi.org/10.1017/nws.2020.3 Published online by Cambridge University Press

interaction. The associated statistic for an actor i given the process $y(t)$ (in this case $s_i^{\text{joining}}(y(t))$ or $s_{i,g}^{\text{leaving}}(y(t))$) is defined by the value of the attribute A_h:

$$s_{\text{ego};i}(y(t)) = a_{i,h}(t) \tag{9}$$

This effect can further be used time-varying, location-varying, or event-varying attributes, to test whether individuals are more or less likely to interact in certain periods, locations, or situations. Members of an organization might, for example, be more likely to interact during lunch breaks, in the cafeteria, or during a corporate event.

Aspiration refers to the tendency of actors to join or leave a group g depending on its members or its own attributes. First, we can define a *size* effect to test whether the size of a group can influence the tendency of actors to join or leave it (steps 2 and 3). This effect is calculated as the in-degree of this group within the interaction network, similarly to *popularity* or *preferential attachment* effects (e.g., Merton, 1968).

$$s_{\text{size},g}(y(t)) = \sum_{j \in \mathscr{I}(t)} x_{j,g}(t) \tag{10}$$

A novelty of this model is its flexibility regarding group attributes. For a given actor attribute h, one can construct various descriptive values of the distribution of this attribute among the members of group g. It is then important to keep such values independent from the size of g, in order to be able to compare interactions in groups of two, three, and so on. We use the function $\bar{f}$ as a placeholder:

$$\bar{f}(g, h, y(t)) = \bar{f}([\![a_{j,h}(t)]\!]_{j \in \mathscr{I}^{(g)}}) \tag{11}$$

that can represent, among others, the mean (e.g., the average age of a group), the minimum, or the maximum (e.g., the lowest or highest status individual in the group). The effect definition for a function $\bar{f}$ and an attribute h is then defined with a *group alter* effect as follows:

$$s_{\text{group-alter},g}(y(t)) = \bar{f}(g, h, y(t)) \tag{12}$$

We can extend this effect to group attributes that are exogenously defined, for instance, if we have some qualitative information on the purpose of this group or the content of the interactions in this group.

Homophily (McPherson et al., 2001) can be generalized to the group context as well if we assume that individuals are more likely to join or stay when a high number of actors are similar to them. It can be operationalized by counting the number of group members j in g with the exact same value $a_{j,h}(t)$ as $a_{i,h}(t)$ by using an indicator function $I(\cdot)$, which we name a *same* effect. This is normalized by the size of g to keep the statistic comparable over the whole model.

$$s_{\text{same},i,g}(y(t)) = \frac{\sum_{j \in \mathscr{I}^{(g)}(t)} I\big[a_{j,h}(t) = a_{i,h}(t)\big]}{|\mathscr{I}^{(g)}(t)|} \tag{13}$$

When the mere presence of an actor with the same attribute as i is relevant, we can also define a *presence same* effect:

$$s_{\text{presence-same},i,g}(y(t)) = \max\left(\big[a_{j,h}(t) = a_{i,h}(t)\big]_{j \in \mathscr{I}^{(g)}(t)}\right) \tag{14}$$

We can also consider a broader definition of *homophily* by looking at the tendency of the actor i to choose groups in which actors' attributes are close to hers. This can be done by calculating the absolute difference between those values, normalized again. This *difference* statistic is null when every single member of the group has the same attribute as actor i and increases as soon as one actor is different:

$$s_{\text{difference},i,g} = \frac{\sum_{j \in \mathscr{I}^{(g)}(t)} |a_{j,h}(t) - a_{i,h}(t)|}{|\mathscr{I}^{(g)}(t)|} \tag{15}$$

When the actor's attribute is comparable to a group measure (e.g., the average of attributes within a group), we can also study the tendency of actors to choose or leave groups who have an average, minimum, or maximum value similar to theirs.

Conformity mechanisms can be constructed by comparing the distribution of group members' attributes to an exogenously defined norm value. Thereby, we can test whether individuals are more likely to join a group where individuals are on average close to a norm (say, an expected level of performance). This highlights how the framework can adapt to various attribute-related mechanisms that may matter in specific empirical contexts.

3.2 Mechanisms related to relational ties

This model can also test the influence of relational variables (e.g., friendship, collaboration) between actors on group interactions. These variables are represented by weighted or binary matrices $Z^{(h)}(t)$ that can vary through time exogenously to the explained process.

The most basic tie-related mechanism expresses the tendency of an individual i to interact with actors j with whom they have a tie (i.e., $z_{i,j}^{(h)}(t) \neq 0$). We can first consider the preference of individuals to choose or leave groups that contain a high proportion of individuals they are related to. We refer to this as a *tie* effect. We can also define a *presence tie* effect if we consider that only the presence of at least one connected actor matters.

$$s_{\text{tie},i,g}(y(t)) = \frac{\sum_{j \in \mathscr{I}^{(g)}(t)} z_{i,j}^{(h)}(t)}{|\mathscr{I}^{(g)}(t)|} \tag{16}$$

$$s_{\text{presence-tie},i,g}(y(t)) = \max \left[z_{i,j}^{(h)}(t) > 0 \right]_{j \in \mathscr{I}^{(g)}(t)} \tag{17}$$

In the case of a weighted network, the *tie* effect can be defined from the weighted sum of the ties and the *presence-tie* effect can be understood as the presence of at least one positive tie (or any other threshold).

Similar to the logic outlined above, we can define transitivity, degree-popularity, or assortativity effects known from statistical network modeling.

3.3 Mechanisms related to previous interactions

The probability of changes in the interaction network might further depend on previous and ongoing interactions. The Markovian framework of the model allows us to store information on past changes as networks or attributes in the process state in Equation (1). Here, we consider, for instance, the network $X^{\text{past}}(t)$ that defines how many times actors have interacted with one another prior to time t. Moreover, additional networks may be included in the process state that represent previous interactions within specific time windows, similar to the approached proposed by Stadtfeld & Block (2017).

All network-related effects introduced above can be transformed to relate to prior interactions by replacing networks $Z^{(h)}(t)$ in the above equations with $X^{\text{past}}(t)$. We believe that accounting for previous interactions, in particular within time windows, is a powerful approach to overcome over-simplistic network models in terms of assumed time homogeneity.

Most importantly, one can then test whether people who have been interacting with group members in the past (or in a given window) are more likely to join their group. We call this the *inertia* effect and define it similar to the *tie* effect above:

$$s_{\text{inertia},i,g}(y(t)) = \frac{\sum_{j \in \mathscr{I}^{(g)}(t)} x_{i,j}^{\text{past}}(t)}{|\mathscr{I}^{(g)}(t)|} \tag{18}$$

$$s_{\text{presence-inertia},i,g} = \max \left[x_{i,j}^{\text{past}} > 0 \right]_{j \in \mathscr{I}^{(g)}(t)} \tag{19}$$

https://doi.org/10.1017/nws.2020.3 Published online by Cambridge University Press

Effects related to the degree of actors can further test whether individuals who have interacted with many others (recently "popular" actors) have a higher tendency to join or leave interactions (*ego popularity* effect) or to be chosen as interaction partners (*alter popularity* effect):

$$s_{\text{ego-popularity},i}(y(t)) = \sum_{j\in\mathscr{I}(t)} x_{i,j}^{(\text{past})}(t) \tag{20}$$

$$s_{\text{alter-popularity},i,g}(y(t)) = \text{mean}\Big[\sum_{k\in\mathscr{I}(t)} x_{j,k}^{(\text{past})}(t)\Big]_{j\in\mathscr{I}^{(g)}(t)} \tag{21}$$

When no time window is defined, we recommend to normalize all these statistics to avoid time heterogeneity in the model.

4. Implementation

4.1 Estimation

The model described in Section 2 contains three sub-parts that are parameterized by the two vectors α and β. The estimation of these parameters is done in a similar way as for the original DyNAM (Stadtfeld & Block, 2017). The likelihood of observing the given data under the rate and choice models defined by these parameters can be calculated as the sum up the partial log-likelihoods derived from the probabilities in the Equations (6) and (8). The values of α and β are then optimized to maximize these two likelihood values, following a standard Newton–Raphson procedure (Deuflhard, 2004).

For more details on this calculation, one can refer to the details provided by Stadtfeld (2012) and Stadtfeld & Block (2017). An important feature of the estimation is that it can be run in parallel over different events. Moreover, the sizes of the actors' opportunity sets within this model are considerably reduced compared to the original DyNAM definitions, which also contributes to greatly decreasing the computational costs and speed up the estimation routine. We thus foresee that this method could scale up to large datasets as collected, although computational limits have not been explored systematically.

4.2 Software

The data preparation and estimation routine are currently implemented within the R package *Goldfish* that also includes the estimation of DyNAM and REM models. The software is available on the website of the Social Networks Lab at ETH Zürich.[1]

4.3 Data requirements

While the mathematical model determines the agency and timing of actors' actions, empirical data can be less detailed. Records of face-to-face interactions as collected via wearable sensors, for example, usually consist of a list of time-stamped dyadic edges that indicate when two individuals were found to be physically close enough to interact. Each observation contains the identifiers of two individuals as well as two time-stamps that indicate when the interaction began and when it stopped. For the presented model, these edges must be translated into events ω, each defined by an actor, an interaction group, a time-stamp, and a variable indicating whether it is a joining or leaving event. Thus, the one-mode network constructed from the original observations is transformed into the two-mode logic of our model. In addition, a list of events indicating the times of formation and dissolution of the second mode nodes is maintained (but these events are not modeled themselves). Figure 2 presents a simple example of this translation for four fictitious actors.

We construct joining events as follows. If a dyadic edge is created between actors A and B and these actors are not interacting with any other actor, we add a joining event from A to the isolate node of B or vice versa. Without any information on agency, the directionality is decided at random. We also remove then the isolate node of the sender of the event from the second mode

https://doi.org/10.1017/nws.2020.3 Published online by Cambridge University Press

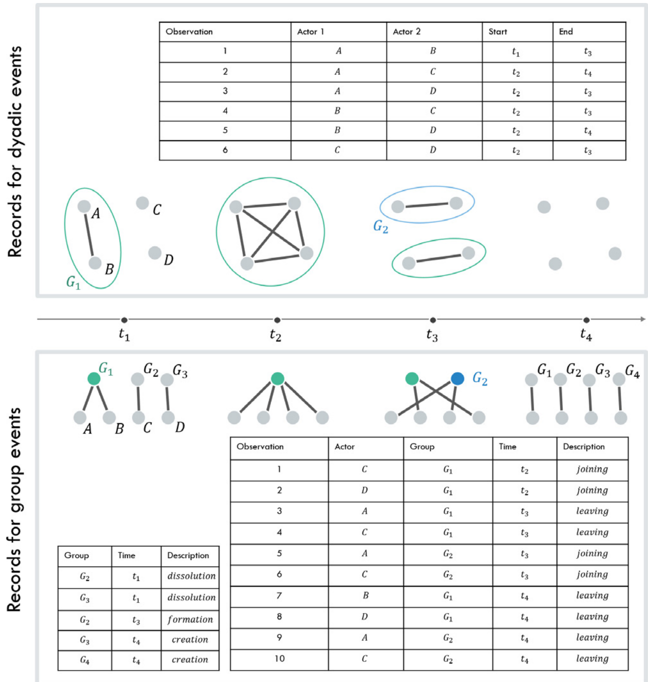

Figure 2. Illustration of the construction of two-mode network events from dyadic interaction records.

set. If an edge is created when A already belongs to a group but B does not, we create a joining event from B to the group of A (and remove the isolate node of B). B is then assumed to interact with all actors present in this group even if some of these edges are not recorded. In cases where both A and B were previously engaged in two different groups, we merge these two groups in a multi-step process. We randomly pick one group to be dissolved and create simultaneous leaving events for the actors of this group and joining events to the group that remains.

Leaving events are created when an edge disappears between two actors A and B. If they are the only members of a group, a leaving event is created from one of the two and an isolate node is created for this actor. If A remains engaged with others, we only create a leaving event for B and an isolate node is created. If one group splits into two subgroups, we randomly pick one of the two and create leaving events for the actors in this subgroup as well as joining events to a new group node. We can represent more complex merges or splits with the same logic.

Some of these practical procedures might affect the model results. Random decisions about the order of actors' actions in the case of simultaneous events influences the rate model's definition and estimation. Randomly picking the groups to create or dissolve also affects the choice model. In cases of splits, for example, only some actors within the group make leaving and joining decisions while others do not. All these effects can be mitigated by performing different randomizations on these procedures and comparing their results.

5. Application 1: the Pizza Party Data set

The following section exemplifies how the proposed model can be used to analyze a data set of precise video records of social interactions between 11 individuals in an informal setting. The data were collected by Elmer et al. (2019) and are available on https://osf.io/rrhxe as well as in the relevant R package *Goldfish*. The presented models describe how mechanisms related to individual attributes, dyadic attributes, and past interactions affect how individuals engaged in their social interactions.

A convenience sample of staff and students from a university were asked to participate in an experiment in exchange for free pizza and beverages. During 76.7 minutes, their social interactions were videotaped and later coded by two independent confederates, with high inter-rater reliability $\kappa = 0.96$ (Cohen, 1968; Landis & Koch, 1977). Participants were mostly males ($N_{\text{male}} = 8$;73%), with a mean age of 29.3 (SD = 3.5). Three were master students, another three were Ph.D. students, four were postdocs, and one was a professor. They belonged to four different work units, and of all pairs of individuals, 26 (47%) knew each other before the event. In total, the data set contains 229 joining and leaving events. Groups that included at least three people were present during 196 (85%) of those events.

5.1 Model specification

To understand the timing of interactions, we jointly estimate the joining and leaving rate models with two different intercepts. To illustrate individual differences, we include in both models the *ego* effects of individuals' age (centered around the mean) and the *ego popularity* effects of their number of past interactions (normalized over actors). For the leaving rate, we also consider parameters taking into account group characteristics. We include a *size* effect accounting for the preference for staying in larger groups. We also use the network of previous acquaintances with a *tie* effect to test whether individuals stay longer with persons they already knew before the event. We finally examine homophilic tendencies: age and seniority are tested with a normalized *difference* effect, while gender and unit affiliation are tested with a normalized *same* effect.

To understand how people chose their groups, we include again the *size* effect, the *tie* effect, the *difference* effects on age and level, and the *same* effects on gender and affiliation. Besides, we add endogenous effects related to previous interactions. First, with an *inertia* effect, we test whether individuals are more likely to interact with others when they have interacted together within the last minute or the last five minutes.[2] Second, with a normalized *alter popularity* effect, we test individuals' preference for individuals who interacted with many others before.

We present here a first model estimated only for attribute effects (Model 1) and a second model that includes structural and endogenous effects of past interactions (Model 2).

5.2 Results

5.2.1 Rate model

The intercepts of the rate model describe the baseline waiting times to either join or leave a group, all other statistics being null. For example, we see in Model 1 (Table 1) that the expected waiting time for an actor of average age to join an interaction is 33 seconds.[3] On the other hand, the waiting times for the leaving rate should be interpreted with caution as they depend on available options for groups (e.g., a fully homophilic group was never possible in the data). However, we can observe that the time to leave an interaction will be much higher than the time to join one, indicating that individuals were engaged in a group most of the time.

In Model 1, the age ego parameter of the joining rate is positive and significant, suggesting that older individuals join social groups slightly faster than younger individuals. The effect of age is also positive but just below the significance level in the leaving rate, indicating that age has a similar but less pronounced effect with regard to leaving a group. Moreover, the age difference to the other group members appears to have a negative effect on leaving an interaction, which points to a preference of individuals to stay in groups with a large age difference. Similarly, the parameter associated with the percentage of individuals of the same unit is positive and significant, which indicates a tendency to stay longer in groups with smaller proportions of direct colleagues. Aside from the effects of age and unit affiliation, we find no evidence for the effects of seniority differences, gender homophily, and previous acquaintances.

Model 2 in Table 1 further includes the effects of group size and the number of past interactions of the focal actors and group members. The parameter for group size in the leaving rate is negative

Table 1. Rate and choice results of the Pizza Party Data set for two models: M1 only contains attribute effects, M2 includes endogenous effects related to the size of the groups and previous interactions

	Model 1			Model 2		
Parameter	est.	s.e.	sig.	est.	s.e.	sig.
Rate model						
Joining parameters						
Intercept joining	−3.507	(0.092)	***	−3.508	(0.093)	***
Age (ego)	0.076	(0.025)	**	0.074	(0.026)	**
Number of past interactions (ego)				−0.028	(0.095)	
Leaving parameters						
Intercept leaving	−6.141	(0.337)	***	−6.109	(0.414)	***
Age (ego)	0.068	(0.037)		0.068	(0.037)	
Difference age (group)	−0.149	(0.072)	*	−0.147	(0.073)	*
Difference seniority (group)	0.270	(0.231)		0.266	(0.236)	
% same gender (group)	0.014	(0.261)		0.021	(0.263)	
% same unit (group)	0.897	(0.379)	*	0.894	(0.394)	*
% known before (group)	−0.228	(0.359)		−0.234	(0.369)	
Size (group)				−0.011	(0.084)	*
Number of past interactions (ego)				0.019	(0.119)	
Log likelihood		−1236			−1236	
AIC		2490			2496	
Choice model						
Age (group)	0.037	(0.017)	*	0.038	(0.018)	*
Difference age (group)	−0.122	(0.070)		−0.113	(0.072)	
Difference seniority (group)	0.105	(0.229)		0.003	(0.235)	
% same gender (group)	0.035	(0.320)		0.220	(0.332)	
% same unit (group)	0.397	(0.425)		0.161	(0.452)	
% known before (group)	1.052	(0.401)	**	1.0455	(0.418)	*
Size (group)				0.017	(0.105)	
Number of past interactions (group)				0.380	(0.182)	*
Repetition 1 min (group)				−0.355	(0.462)	
Repetition 5 min (group)				0.082	(0.239)	
Log likelihood		−182			−178	
AIC		376			375	

***: < 0.001, **: < 0.01, *: < 0.05

and significant, showing that individuals tended to remain longer engaged in larger groups. Both effects for past interactions are non-significant, showing no evidence for endogenous effects in this model. Other parameters remain similar to the ones in Model 1, and the total likelihood is not affected. The small decrease of the AIC points to a very marginal influence of structural and endogenous effects.

5.2.2 Choice model

In the choice model of Model 1 (Table 1), we observe a significant effect of the average age of a group, with a tendency of individuals to choose older groups. Another significant parameter is found for the dyadic effect of previous acquaintances, providing some evidence for individuals'

https://doi.org/10.1017/nws.2020.3 Published online by Cambridge University Press

preference for groups in which they already know more group members. Other effects of age, seniority, and gender homophily are non-significant.

Of all effects added in Model 2 (Table 1), only the average number of past interactions of group members is associated with a significant parameter. This result suggests a popularity effect, with individuals choosing to interact with others who were involved in a higher number of interactions. The size parameter is positive, indicating a preference for larger groups again, although it remains below the significance level. The two other effects are windowed inertia effects for windows of 1 and 5 minutes. Although both effects are not significant, their different directions would suggest an interesting pattern. They would indeed indicate that people chose groups with individuals with whom they had interacted recently (within 5 minutes), but less so if this interaction was very recent (within one minute). Other parameters remain similar, and we can observe a small improvement of the log-likelihood and the AIC.

As explained in Section 4, we applied different randomizations to the data processing step and found that all previous results were robust. We only observed some small variations in the significance level of the significant parameters. Moreover, we solely analyze here the direction and significance levels of our estimates. An in-depth interpretation of effect sizes would require to take into account the opportunity sets of each event carefully.

6. Application 2: the Badge Data set

This section aims at showing how micro temporal patterns of face-to-face interactions in an office setting can be explained by organizational and spatial attributes, as well as endogenous mechanisms related to past interactions. We use the *Badge Data set* collected by Olguín et al. (2008) and further analyzed by Wu et al. (2008) and Dong et al. (2012). This data set contains records of interactions between 37 employees of an IT firm for one month. These interactions were measured by giving each participant a sociometric badge (Olguín et al., 2008) that could detect proximity to other employees within a range of one meter. Each badge also recorded the approximate location, the movements, and the voice of its wearer. The data set further provides information on employees' tasks during the time of the study, as well as a detailed floorplan of the office space.

The sample comprises 37 employees grouped into three branches, 26 of them working as configuration experts, 7 of them being in charge of pricing, and 4 having a coordinator role. Four of them held a managing position. All desks were located on the same floor, close to each other (mean distance = 24.75 m, SD = 3.5), and 40 pairs of employees had their cubicles adjacent (around 6% of the total 666 pairs). Finally, 14 employees' desks were situated along the central corridor that gave access to the coffee machine, the kitchen, and the printers.

We model the interactions recorded during the second week of the study on each day from Monday to Friday.[4] We preprocessed these dyadic events by merging subsequent events occurring within 30 seconds, similarly to the approach suggested by Elmer et al. (2019). This time interval was chosen because shorter intervals created many repeated events, while longer intervals produced a more stable number of events. In total, our data contain 688 joining events and 987 leaving events. Groups with more than two actors account for approximately 3% of those events. Employees' task records are also included in our analyses, with a total of 290 tasks assigned during the week.

6.1 Model specification

Our models are specified according to observations made in the original studies of this data set by Wu et al. (2008) and Dong et al. (2012), as well as other studies examining the link between office layouts and social interactions (Sailer & McCulloh, 2012; Sailer et al., 2012; Wineman et al., 2014).

We first investigate the impact of tasks and organizational roles on interaction patterns. Wu et al. (2008) and Dong et al. (2012) observe that communication was essential for employees to complete their tasks. This finding suggests that the more tasks a person has, the more likely she will interact with others, and perhaps stay longer in these interactions. We test these tendencies with an *ego* effect of the number of ongoing tasks on the rate of joining and leaving interactions. Dong et al. (2012) also note that many interactions occurred among configuration employees, and between configuration and pricing employees. Other studies also find more frequent interactions inside work units (Sailer & McCulloh, 2012; Wineman et al., 2014). We examine the tendencies in the choice and the leaving models with a normalized homophily effect (i.e., *same* effect) on the employees' role and an attraction effect (i.e., *alter* effect) to groups that contain at least one configuration employee. We further explore differences in the interaction behaviors of different branches with an *ego* effect for configuration employees in the rate models.

Second, we investigate the effect of the spatial layout. Dong et al. (2012) note that employees with close desks spent more time together, consistently with other studies (Sailer & McCulloh, 2012; Sailer et al., 2012; Wineman et al., 2014). This effect of proximity is modeled with a *tie* effect of two dyadic attributes. The first is the euclidean distance between two employees' desks, and the second is an indicator for employees to have neighboring desks. Previous research also finds that coffee machines, kitchens, and printers were crucial areas in an office (Sailer et al., 2012; Wineman et al., 2014). We thus include *ego* and *alter* effects for employees having their desk adjacent to the central corridor, as they could have been more frequently visited because of their central position.

Third, we control for differences in group sizes and temporal patterns. We add a *size* effect that represents individuals' tendency to join and stay in groups with more than two individuals. Following the observation of Dong et al. (2012) that some individuals were more often involved in interactions than others, we also include *ego popularity* and *alter popularity* effects. Finally, we explore the propensity of actors to interact if they have recently interacted by using an *ego popularity* effect with a window of an hour in the joining model.[5] We also add an *inertia* effect for windows of 1 hour and 1 day[5] to account for individuals' preference to repeatedly interact with the same individuals within those intervals.

We report the results of a first model with only organizational and spatial effects (Model 1) and another model containing all effects, including size and temporal effects (Model 2).

6.2 Results

6.2.1 Rate model

Results of Model 1 are reported in Table 2. The intercepts of the joining and the leaving models considerably differ, suggesting that individuals waited longer to join an interaction than to leave. This is consistent with the high proportion of short interactions and long times being isolated that we observe in the data. Regarding individual attributes, we observe a positive and significant effect of the number of tasks on the rate of joining and leaving interactions. This indicates that employees with ongoing tasks interacted more often with their colleagues but for shorter times. Roles also seem to have an impact on employees' activities, with configuration employees being faster to leave an interaction or to be left by others. This finding supports the idea that the branches behaved differently and suggests that configuration employees were involved in shorter interactions. We find evidence for role homophily; individuals tend to remain longer with employees of the same branch. Finally, the central position of an employee does not have a significant impact on her joining an interaction but does increase her tendency to stay longer in an interaction. Similarly, individuals stayed longer with central employees. We also observe that interactions with proximate others tended to be shorter.

Model 2 (Table 2) further includes the effects of group sizes and past interactions. We first observe a drop in the values of the log-likelihood, which suggests that the added variables are particularly relevant for explaining the patterns observed in the data. Most of the parameters remain similar to the ones observed in Model 1. An exception for this is the effect related to working in the

https://doi.org/10.1017/nws.2020.3 Published online by Cambridge University Press

Table 2. Rate and choice results for the Badge Data set: M1 only contains attribute effects, M2 includes endogenous effects related to the size of the groups and previous interactions

	Model 1			Model 2		
Parameter	est.	s.e.	sig.	est.	s.e.	sig.
Rate model						
Joining parameters						
Intercept joining	−9.976	(0.089)	***	−10.406	(0.106)	***
Configuration role (ego)	−0.054	(0.085)		−0.353	(0.096)	***
Corridor desk (ego)	0.035	(0.078)		0.095	(0.084)	
N current tasks (ego)	0.056	(0.025)	*	0.054	(0.031)	***
N past interactions (ego)				0.457	(0.029)	***
N past interactions 1 hour (ego)				0.082	(0.002)	***
Leaving parameters						
Intercept leaving	−2.369	(0.301)	***	−2.779	(0.246)	***
Configuration role (ego)	0.551	(0.100)	***	0.418	(0.103)	***
Corridor desk (ego)	−1.554	(0.100)	***	−1.595	(0.100)	***
N current tasks (ego)	0.099	(0.029)	**	0.048	(0.027)	
Configuration role (group)	0.737	(0.106)	***	0.842	(0.106)	***
% same role (group)	−1.744	(0.210)	***	−1.187	(0.178)	***
Corridor desk (group)	−1.430	(0.102)	***	−1.307	(0.097)	***
Distance desks (group)	−0.002	(0.006)		0.015	(0.005)	**
Neighbor desks (group)	0.724	(0.200)	***	0.395	(0.193)	*
Size (group)				2.528	(0.201)	***
N past interactions (group)				-0.169	(0.027)	***
Log likelihood		−10864			−9518	
AIC		21754			19071	
Choice model						
Configuration role (group)	−0.396	(0.092)	***	−0.438	(0.154)	**
Corridor desk (group)	0.141	(0.077)	***	−0.212	(0.123)	
% same role (group)	1.054	(0.109)	***	1.036	(0.181)	***
Distance desks (group)	0.004	(0.003)		−0.008	(0.006)	
Neighbor desks (group)	−0.399	(0.171)	*	0.470	(0.196)	*
Size (group)				4.609	(1.227)	***
N past interactions (group)				0.409	(0.042)	***
Repetition 1 hour (group)				0.490	(0.038)	***
Repetition 24 hours (group)				−0.025	(0.013)	
Log likelihood		−2428			−976	
AIC		4867			1969	

***: < 0.001, **: < 0.01, *: < 0.05

configuration branch, which becomes significant and negative in the joining rate. This result indicates that configuration employees joined interactions at a lower rate than others. In the leaving rate, the effect of having a task disappears, while the distance between desks becomes a positive predictor for the rate of leaving. This second result is in line with the common finding that individuals tend to spend more time with others sitting close to them. Regarding new effects, we observe a positive and significant parameter for group sizes, indicating that individuals stay longer in dyadic interactions. We also observe significant parameters for the popularity effects, showing

https://doi.org/10.1017/nws.2020.3 Published online by Cambridge University Press

that individuals who had already interacted more were more likely to join an interaction and that others were more likely to stay with them too. The effect of previous interactions with a window of an hour shows a similar trend, which indicates that recent interactions play a particular role in interaction processes.

6.2.2 Choice model

Table 2 shows the results of the choice models. Model 1 suggests that employees were more likely to interact with others of the same branch and less with configuration employees. This finding complements the previous result that individuals stayed longer with others from the same branch but left faster configuration employees. Results concerning spatial attributes show that having neighboring desks is a negative predictor for social interactions, while the parameter associated with desk distances remains non-significant. The effect of having a central desk is significant and positive, suggesting that individuals occupying a desk along the corridor were more often chosen as interaction patterns. This supports the idea that individuals who were more central in the office were interacting more with others.

Results for Model 2 show that some parameters change when including the effect of group sizes and endogenous mechanisms. Although the findings related to employees' roles remain unchanged, the conclusions that can be drawn for the spatial effects differ. The parameter related to neighboring desks becomes positive, and the effect of having a central desk disappears. Thus, the results of Model 2 suggest that neighbors were more likely to interact, which provides some evidence for the impact of co-location on social interactions that previous studies found. They do not show, however, that individuals who were more central in the office were interacting more with others. Regarding new effects, we first observe a significant propensity to choose groups with more than two individuals when they are available. Moreover, individuals who interacted many times before seem to be chosen more often, which is coherent with the popularity effects observed in the rate models. Finally, having interacted in the last hour is a significant predictor for the choice of the interaction partner. The same effect applied to a window of 1 day is not significant. Overall, we observe a substantial decrease of the log-likelihood and the AIC in Model 2. This drop suggests that these endogenous effects are important to explain the patterns observed in these data and that Model 1 is likely to be misspecified. It is important to note that some endogenous effects might reflect the effect of missing covariates that were not available in the data set. For example, previous studies find that shared projects and gender homophily also predict interactions among office mates (Potter et al., 2015), which could be captured by the *inertia* effects.

7. Discussion and conclusions

Face-to-face interactions are a central aspect of our social lives. Many of these interactions occur during social occasions in which conversations not only unfold in dyads but also in groups. Such group interactions naturally emerge in various settings, for example, in the family, at school, at the workplace, and during occasions such as social gatherings or scientific conferences. The empirical study of these group interactions promises to open insights into social mechanisms proposed by many theories from sociology, social psychology, or management literature. The analysis of data collected on face-to-face interactions, however, remains difficult. Current statistical methods are not suited for the group structures typically present in such interactions and can only model mechanisms operating at a dyadic level. New data collection strategies to measure interaction dynamics on a fine-grained level have recently been proposed, but the challenge remains to align them with statistical methodology.

This article proposes a new statistical model, DyNAM-i, that addresses the challenges of modeling the complex dynamics of face-to-face interactions. It extends the class of relational event models in three ways. First, this model is suited for group interaction data and allows us to analyze the formation and dissolution of groups without any restriction on their size. Second, it can incorporate complex interdependencies between interaction events by extending the common social network mechanisms—such as inertia, transitivity, popularity, and homophily—to the settings of groups. The DyNAM-i model can be flexibly specified and be used to test how aggregate outcomes, such as group sizes and group compositions, can affect how individuals engage in interactions. Third, it is built to align with newly available data collection strategies in the context of group interactions. It can thus be applied to dyadic time-stamped events as usually collected by social sensors, video data, or RFID-, Bluetooth-, or wifi-based proximity measures (Cattuto et al., 2010; Elmer et al., 2019; Hong et al., 2016; Olguín et al., 2008; Pentland, 2008; Sapiezynski et al., 2015). Other data of the same format of which the underlying processes align with the assumptions of this model could theoretically also be analyzed (e.g., time-stamped data on the joining and leaving of employees of different companies). A software implementation of the estimation method is available in the R package *Goldfish*. It is worth noting that the computational costs of the model estimation are lower than most other network models, which allows the analysis of comparatively large data sets.

In two different applications, we demonstrate that this model can provide insights on micro-temporal patterns of social interactions. Our first example setting is a small social gathering during which individuals lively engaged in small groups interactions (Elmer et al., 2019). Our models show that interactions in this context could be explained by age effects (older individuals joined and left interactions more often, and individuals remained longer in groups with higher age differences), by previous acquaintances (individuals joined groups where they knew more members), and group affiliation (individuals interacted longer with members of other research groups). In the second application, we treat a large data set of interaction records between employees of a same office (Dong et al., 2012; Olguín et al., 2008; Wu et al., 2008). In this context, face-to-face interactions were mainly dyadic and punctual. Our analyses suggest that these interactions were mostly driven by employees' organizational roles (we found evidence for role homophily and preference for certain roles), spatial layouts (employees were more likely to interact with their neighbors), and endogenous effects of past interactions (individuals tended to repeat previous interactions, and some employees appeared more popular than others).

Some limitations remain. One can be derived from its actor-oriented nature: Individuals are assumed to be in control of their actions and to "optimize" their own interaction situation by joining and leaving groups. This assumption might be problematic in settings where individuals coordinate their interactions, and the agency is situated at the group level. Further, complex changes of the interacting groups such as splits and mergers might not be well represented by sequences of individual actions to join and leave groups. Finally, a practical challenge in applying the model is the collection of data that accurately represents social behavior and the design of appropriate data preprocessing strategies.

Future developments should address some of these limitations. In particular, the mechanisms underlying actors' coordination, or groups splitting and merging, could be circumvented by developing models similar to tie-oriented network models in which the agency assumptions are relaxed. Furthermore, we highlighted that group interactions might be related to crucial individual, relational, and group-level outcomes. Emerging interaction patterns may, for example, affect individuals' well-being, who they become friends with, or whether a team is solving problems successfully. Understanding such dynamics would require to extend the model so that dynamic-dependent variables on these levels can be modeled simultaneously. Finally, some crucial observations, such as agency in group mergers or detailed spatial information, are difficult to collect with current techniques. We hope that the availability of the new model will inspire further data collection strategies that aim at closing these empirical gaps.

We believe that the possible applications of the DyNAM-i model are manifold. It could be used to study, for example, the interaction routines of families, the dynamics of children groups at school, the communication structure of an organization, or the emergence of new social groups at social gatherings or conferences. We argued that the dynamics of face-to-face interactions in such social occasions are theoretically meaningful and a core element of social lives. The new model aims at contributing to closing the current gap between theory and empirical research on the dynamics of face-to-face interactions in social groups.

Acknowledgments. We thank members of the Social Networks Lab at ETH Zürich for helpful comments and suggestions. We also thank Krishna Chaitanya and Prateek Purwar for their help in collecting the Pizza Party Dataset as well the participants of the study. The study was supported by the Swiss National Science Foundation with the "Digital Lives" Grant 10DL17_183008. Per Block is supported by the Leverhulme Centre for Demographic Science, The Leverhulme Trust.

Conflict of interest. The authors report no conflicts of interest.

Notes

1 www.sn.ethz.ch/research/goldfish; An updated version that includes a documentation of DyNAM-i will be available by Summer 2020.
2 The choice of the time windows is arbitrary. In substantive articles with specific research questions, the choice of these lengths should be argued for and tested with robustness analyses.
3 Calculated as $\frac{1}{\exp(-3.507+0*0.076)}$ seconds.
4 To account for the difference between days and nights, we estimate our rate models assuming that 07:00 in the morning immediately followed 21:00 of the previous day.
5 The lengths of time windows are arbitrarily chosen here, but alternative specifications could be used.

References

Allen, J., James, A. D., & Gamlen, P. (2007). Formal versus informal knowledge networks in r&d: A case study using social network analysis. *R&d Management*, *37*(3), 179–196.
Amati, V., Lomi, A., & Mascia, D. (2019). Some days are better than others: Examining time-specific variation in the structuring of interorganizational relations. *Social Networks*, *57*, 18–33.
Bales, R. F. (1950). *Interaction process analysis; a method for the study of small groups*. Addison-Wesley.
Block, P. (2018). Network evolution and social situations. *Sociological Science*, *5*, 402–431.
Brandenberger, L. (2019). Predicting network events to assess goodness of fit of relational event models. *Political Analysis*, *27*(4), 556–571.
Brandes, U., Lerner, J., & Snijders, T. A. (2009). Networks evolving step by step: Statistical analysis of dyadic event data. In *2009 international conference on advances in Social network analysis and mining* (pp. 200–205). IEEE.
Butts, C. T. (2008). A relational event framework for social action. *Sociological Methodology*, *38*(1), 155–200.
Carley, K. M., & Krackhardt, D. (1996). Cognitive inconsistencies and non-symmetric friendship. *Social Networks*, *18*(1), 1–27.
Cattuto, C., den Broeck, W., Barrat, A., Colizza, V., Pinton, J.-F., & Vespignani, A. (2010). Dynamics of person-to-person interactions from distributed RFID sensor networks. *PloS One*, *5*(7), e11596.
Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin*, *70*(4), 213.
Collins, R. (2014). *Interaction ritual chains*, vol. 62, Princeton university press.
Deuflhard, P. (2004). *Newton methods for nonlinear problems*, volume 35 of springer series in computational mathematics.
Dong, W., Olguin-Olguin, D., Waber, B., Kim, T., & Pentland, P. (2012). Mapping organizational dynamics with body sensor networks. In *2012 Ninth International Conference on Wearable and Implantable Body Sensor Networks* (pp. 130–135). IEEE.
Elmer, T., Chaitanya, K., Purwar, P., & Stadtfeld, C. (2019). The validity of RFID badges measuring face-to-face interactions. *Behavior Research Methods*, *51*(5), 2120–2138.
Feld, S. L. (1982). Social structural determinants of similarity among associates. *American Sociological Review*, *47*, 797–801.
Fischer, C. S. (1982). *To dwell among friends: Personal networks in town and city*. Chicago: University of Chicago Press.
Goffman, A. (2019). Go to more parties? Social occasions as home to unexpected Turning points in life trajectories. *Social Psychology Quarterly*, *82*(1), 51–74
Goffman, E. (1967). *Interaction ritual: Essays on face-to-face interaction.* Oxford, England: Aldine.
Holland, P. W., & Leinhardt, S. (1977). A dynamic model for social networks. *Journal of Mathematical Sociology*, *5*(1), 5–20.
Hong, H., Luo, C., & Chan, M. C. (2016). Socialprobe: Understanding social interaction through passive wifi monitoring. In *Proceedings of the 13th international conference on mobile and Ubiquitous systems: Computing, networking and services* (pp. 94–103). ACM.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174.
Lazarsfeld, P. F., & Merton, R. K. (1954). Friendship as a social process: A substantive and methodological analysis. *Freedom and Control in Modern Society*, *18*(1), 18–66.
Lazega, E., Lemercier, C., & Mounier, U. (2006). A spinning top model of formal organization and informal behavior: Dynamics of advice networks among judges in a commercial court. *European Management Review*, *3*(2), 113–122.
Lerner, J., & Lomi, A. (2019). Reliability of relational event model estimates under sampling: How to fit a relational event model to 360 million dyadic events. arXiv preprint arXiv:1905.00630.
Lewin, K. (1947). Frontiers in group dynamics: Concept, method and reality in social science; social equilibria and social change. *Human Relations*, *1*(1), 5–41.
Lusher, D., Koskinen, J., & Robins, G. (Eds.) (2013). *Exponential random graph models for social networks*. New York: Cambridge University Press.
Marcum, C. S., & Butts, C. T. (2015). Constructing and modifying sequence statistics for relevent using informr in r. *Journal of Statistical Software*, *64*(5), 1–36.
McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics*, chapter 4 (pp. 105–142). New York: Academic Press Inc.
McPherson, J. M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*, 415–444.
Merton, R. K. (1968). The Matthew effect in science. *Science*, *159*(3810), 56–63.
Moody, J. (2001). Race, school integration, and friendship segregation in america. *American Journal of Sociology*, *107*(3), 679–716.
Mulder, J., & Leenders, R. T. A. (2019). Modeling the evolution of interaction behavior in social networks: A dynamic relational event approach for real-time analysis. *Chaos, Solitons & Fractals*, *119*, 73–85.
Olguín, D. O., Waber, B. N., Kim, T., Mohan, A., Ara, K., & Pentland, A. (2008). Sensible organizations: Technology and methodology for automatically measuring organizational behavior. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *39*(1), 43–55.
Pattison, P., & Robins, G. (2002). 9. neighborhood-based models for social networks. *Sociological Methodology*, *32*(1), 301–337.
Pentland, A. S. (2008). *Honest signals. How hey shape our world*. Cambridge, Massachussetts: The MIT Press.
Perry, P. O., & Wolfe, P. J. (2013). Point process modelling for directed interaction networks. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *75*(5), 821–849.
Potter, G. E., Smieszek, T., & Sailer, K. (2015). Modeling workplace contact networks: The effects of organizational structure, architecture, and reporting errors on epidemic predictions. *Network Science*, *3*(3), 298–325.
Robins, G. (2015). *Doing social network research*. London: SAGE.
Sailer, K., & McCulloh, I. (2012). Social networks and spatial configuration—how office layouts drive social interaction. *Social Networks*, *34*(1), 47–58.
Sailer, K., Marmot, A., & Penn, A. (2012). Spatial configuration, organisational change and academic networks. Paper presented at the Conference on Applied Social Network Analysis, 4–7 September 2012, Zürich, Switzerland.
Sapiezynski, P., Stopczynski, A., Gatej, R., & Lehmann, S. (2015). Tracking human mobility using wifi signals. *PloS One*, *10*(7), e0130824.
Schecter, A., Pilny, A., Leung, A., Poole, M. S., & Contractor, N. (2018). Step by step: Capturing the dynamics of work team process through relational event sequences. *Journal of Organizational Behavior*, *39*(9), 1163–1181.
Schroeder, H., & Lovell, H. (2012). The role of non-nation-state actors and side events in the international climate negotiations. *Climate Policy*, *12*(1), 23–37.
Sekara, V., Stopczynski, A., & Lehmann, S. (2016). Fundamental structures of dynamic social networks. *Proceedings of the National Academy of Sciences*, *113*(36), 9977–9982.
Simmel, G., & Hughes, E. C. (1949). The sociology of sociability. *American Journal of Sociology*, *55*(3), 254–261.
Snijders, T. A. (1996). Stochastic actor-oriented models for network change. *Journal of Mathematical Sociology*, *21*(1–2), 149–172.
Snijders, T. A. (2017). Stochastic actor-oriented models for network dynamics. *Annual Review of Statistics and its Application*, *4*, 343–363.
Snijders, T. A. & Lomi, A. (2019). Beyond homophily: Incorporating actor variables in statistical network models. *Network Science*, *7*(1), 1–19.
Snijders, T. A., Van de Bunt, G. G., & Steglich, C. E. (2010). Introduction to stochastic actor-based models for network dynamics. *Social Networks*, *32*(1), 44–60.
Spagnola, M., & Fiese, B. H. (2007). Family routines and rituals: A context for development in the lives of young children. *Infants & Young Children*, *20*(4), 284–299.
Stadtfeld, C. (2012). *Events in social networks. A stochastic actor-oriented framework for dynamic event processes in social networks*. Karlsruhe: KIT Scientific Publishing.

Stadtfeld, C., & Block, P. (2017). Interactions, actors, and time: Dynamic network actor models for relational events. *Sociological Science*, *4*(1), 318–352.

Stadtfeld, C., & Geyer-Schulz, A. (2011). Analyzing event stream dynamics in two-mode networks: An exploratory analysis of private communication in a question and answer community. *Social Networks*, *33*(4), 258–272.

Stadtfeld, C., Hollway, J., & Block, P. (2017). Dynamic network actor models: Investigating coordination ties through time. *Sociological Methodology*, *47*(1), 1–40.

Tajfel, H., Turner, J. C. (1979). An integrative theory of intergroup conflict. In W.G. Austin & S. Worchel (Eds.), *The Social Psychology of Intergroup Relations*. Monterey: Brooks-Cole.

Tajfel, M. (1970). Experiments in intergroup discrimination. *Scientific American*, *223*, 96–102.

Verbrugge, L. M. (1977). The structure of adult friendship choices. *Social Forces*, *56*(2), 576–597.

Vu, D., Pattison, P., & Robins, G. (2015). Relational event models for social learning in moocs. *Social Networks*, *43*, 121–135.

Wang, W., Bai, X., Xia, F., Bekele, T. M., Su, X., & Tolba, A. (2017). From triadic closure to conference closure: The role of academic conferences in promoting scientific collaborations. *Scientometrics*, *113*(1), 177–193.

Wasserman, S. (1980). Analyzing social networks as stochastic processes. *Journal of the American Statistical Association*, *75*(370), 280–294.

Wineman, J., Hwang, Y., Kabo, F., Owen-Smith, J., & Davis, G. F. (2014). Spatial layout, social structure, and innovation in organizations. *Environment and Planning B: Planning and Design*, *41*(6), 1100–1112.

Wu, L., Waber, B. N., Aral, S., Brynjolfsson, E., & Pentland, A. (2008). Mining face-to-face interaction networks using sociometric badges: Predicting productivity in an it configuration task. Available at SSRN 1130251.

Wynn, J. R. (2016). On the sociology of occasions. *Sociological Theory*, *34*(3), 276–286.

Cite this article: Hoffman M., Block P., Elmer T., and Stadtfeld C. (2020). A model for the dynamics of face-to-face interactions in social groups. *Network Science* **8**, S4–S25. https://doi.org/10.1017/nws.2020.3